

3

The development of the PISA context questionnaires

| | |
|---|----|
| Overview..... | 50 |
| The conceptual structure..... | 51 |
| ▪ A conceptual framework for PISA 2006 | 51 |
| Research areas in PISA 2006 | 55 |
| The development of the context questionnaires..... | 57 |
| The coverage of the questionnaire material..... | 58 |
| ▪ Student questionnaire..... | 58 |
| ▪ School questionnaire..... | 59 |
| ▪ International options..... | 59 |
| ▪ National questionnaire material..... | 60 |
| The implementation of the context questionnaires..... | 60 |



OVERVIEW

In its Call for Tender for PISA 2006, the PISA Governing Board (PGB) established the main policy issues it sought to address in the third cycle of PISA. In particular, the PGB required PISA 2006 to collect a set of basic demographic data as a core component that replicated key questions from the previous cycles. In addition, PISA 2006 needed to address issues related to important aspects of students' attitudes regarding science, information about students' experience with science in and out of school, motivation for, interest in and concern about science, and engagement with science-related activities.

Since the impact of out-of-school factors was considered of particular interest in a cycle where science was the major domain, the PGB recommended the inclusion of a parent questionnaire as an optional instrument, in order to collect additional information on issues such as science-related parental expectations and attitudes, as well as possible family investment in activities aimed at developing students' interest and learning in scientific areas.

The PISA 2006 Project consortium undertook the operationalisation of these goals with the assistance of a variety of experts. In particular, a Questionnaire Expert Group (QEG) was established, consisting of experts from a variety of research backgrounds and countries (see Appendix 8). The consortium and the QEG worked together to develop the contextual framework for PISA 2006 and the contextual instruments. Other experts were consulted where appropriate, especially some members of the Science Expert Group.

An initial step was the development of an organising conceptual structure which allowed the mapping of the PGB's priority policy issues to the design of PISA 2006. One important objective of the conceptual structure was to facilitate the development and choice of research areas that combine policy relevance effectively with the strengths of the PISA design. To aid this, a set of criteria established by the INES (International Indicators of Educational Systems) Network A were used:

- First, the research area must be of enduring policy relevance and interest. That is, a research area should have policy relevance, capture policy-makers' attention, address their needs for data about the performance of their educational systems, be timely, and focus on what improves or explains the outcomes of education. Further, a research area should be of interest to the public, since it is this public to which educators and policy-makers are accountable;
- Second, research areas must provide an internationally comparative perspective and promise significant added value to what can be accomplished through national evaluation and analysis. This implies that research areas need to be both relevant (*i.e.* of importance) and valid (*i.e.* of similar meaning) across countries;
- Third, there must be some consistency in the approach of each research area with PISA 2000 and PISA 2003;
- Fourth, it must be technically feasible and appropriate to address the issues within the context of the PISA design. That is, the collection of data about a subject must be technically feasible in terms of methodological rigour and the time and costs (including opportunity costs) associated with data collection.

The resulting research areas are listed below and described in more detail later in the chapter:

- Student's engagement in science
- Science attainment and the labour market
- Teaching and learning science
- Scientific literacy and environment
- Organisation of educational systems



THE CONCEPTUAL STRUCTURE

A conceptual framework for PISA 2006

Both the basic criteria for developing a conceptual framework presented above, and more comprehensive reviews of educational models (Scheerens and Bosker, 1997) reveal the complexity of variables and relationships that potentially influence student outcomes. The field is at the crossroads between a number of sociological, psychological, and cognitive theories, which all contribute important components to the overall picture.

Developing a new single, encompassing educational model for PISA would add little value to the many models already available in the literature. Rather than imposing unnecessary theoretical constraints on the thematic analyses that will be conducted using the study database, the primary role of the PISA conceptual structure for questionnaire development was to map the many components of existing models, to ensure that none of the essential dimensions are omitted from the data collection. These components were then checked against the general framework used for the OECD education indicators (INES) and the PGB priorities for PISA 2006.

This mapping also facilitated discussions around the feasibility and appropriateness of implementation within the constraints of the PISA design. In particular, the following aspects were considered, both in terms of restrictions and of potentialities related to the study design:

- PISA measures knowledge and skills for life and so does not have a strong curricular focus. This limits the extent to which the study is able to explore relationships between differences in achievement and differences in the implemented curricula. On the other hand, consideration was given to the out-of-school factors with a potential of enhancing cognitive and affective learning outcomes;
- PISA students are randomly sampled within schools, not from intact classrooms or courses and therefore come from different learning environments with different teachers and, possibly, different levels of instruction. Consequently, classroom-level information could only be collected either at the individual student level or at the school level;
- PISA uses an age-based definition of the target population. This is particularly appropriate for a yield-oriented study, and provides a basis for in-depth exploration of important policy issues, such as the effects of a number of structural characteristics of educational systems (e.g. the use of comprehensive vs. tracked study programmes, or the use of grade repetition). On the other hand, the inclusion in the study of an increasing number of non-OECD countries (where the enrolment rate for the 15-year-olds age group is maybe less than 100%) requires that retention be taken into account in the analysis of between-countries differences;
- The cross-sectional design in PISA does not allow any direct analysis of school effects over time. However, the cyclic nature of the study will permit not only the investigation of change in the criterion measures, but also in the effects of rates of change in the predictor variables.

Many conceptual models to explain learning outcomes distinguish different levels that relate both to the entities from which data might be collected and to the multi-level structure of national education systems (Scheerens 1990). Four levels can be distinguished:

- The education system as a whole (setting the context for teaching and learning);
- The educational institutions (schools but also other providers of education);
- The instructional setting and the learning environment within the institutions (classrooms, courses);
- The individual participants in learning activities (students).



A second dimension commonly found in many conceptual models groups the indicators at each of the above levels further into the following categories:

- Antecedents are those factors that affect policies and the way instruction is organised, delivered and received. It should be noted that they are usually specific for a given level of the education system and that antecedents at a lower level of the system may well be policy levers at a higher level (e.g. for teachers and students in a school, teacher qualifications are a given constraint while, at the level of the education system professional development of teachers is a key policy lever);
- Processes group information on the policy levers or circumstances that shape the outputs and outcomes at each level;
- Indicators on observed outcomes of education systems, as well as indicators related to the impact of knowledge and skills for individuals, societies and economies, are grouped under outcomes.

The four levels and the three aspects can be visualised as a two-dimensional grid with 12 potential variable types (Figure 3.1). This basic conceptualisation has been adapted from the conceptual framework for the Second IEA Study of Mathematics (Travers and Westbury, 1989; Travers, Garden and Rosier, 1989) and also provided a conceptual basis for the planning of context questionnaires for the first two PISA surveys (Harvey-Beavis, 2002; OECD, 2005). As noted earlier, data on the instructional settings can only be collected at the individual or institutional level. However, conceptually they are still related to the level of the instructional settings (classroom, courses).

Figure 3.1 shows the basic components of this two-dimensional grid. It consists of four levels and variables at each level are classified as antecedents, processes or outcomes:

- At the system-level, the macroeconomic, social, cultural and political context sets constraints for the educational policies in a country. Outcomes at the system-level are not only aggregated learning outcomes but also equity-related outcomes;
- At the level of the educational institution, characteristics of the educational provider and its community context are antecedents for the policies and practices at the institutional level as well as the school climate for learning. Outcomes at this level are aggregates of individual learning outcomes and also differences in learning outcomes between sub-groups of students;

Figure 3.1

Conceptual grid of variable types

| Antecedents | Processes | Outcomes |
|--|--|--|
| Level of the educational system | | |
| Macro-economic, social, cultural and political context | Policies and organisation of education | Outcomes at the system level |
| Characteristics of educational institutions | Institutional policies and practice | Outcomes at the institutional level |
| Level of instructional units | | |
| Characteristics of instructional units | Learning environment | Outcomes at the level of instructional units |
| Level of individual learners | | |
| Student background and characteristics | Learning at the individual level | Individual learning outcomes |



- At the level of the instructional units, characteristics of teachers and the classrooms/courses are antecedents for the instructional settings and the learning environment; learning outcomes are aggregated individual outcomes;
- At the student level, characteristics (like gender, age, grade) and background (like social status, parental involvement, language spoken at home) are antecedents for the individual learning process and learning outcomes (both cognitive and affective).

It should be noted that learning outcome variables consist not only of cognitive achievement but also of other potential learning outcomes. These include self-related cognitions (self-concept, self-efficacy), long-term interest in a subject or domain, educational expectations and aspirations as well as social outcomes like well-being and life skills.

While this mapping is useful for planning the coverage of the PISA questionnaires it is also important to supplement it with recognition of the dynamic elements of the educational system. System-level variables are important when interpreting relationships between variables at the lower levels and contradictory findings across countries are often due to differences in the structure of the educational systems.

From the existing conceptual frameworks and subsequent research one can derive hypotheses about (at least some of) the relationships between the elements in this two-dimensional grid. Typically, existing conceptual models assume antecedents to influence processes, which in turn produce learning outcomes, and conditions on higher levels are usually supposed to impact on those at lower levels (Scheerens, 1990).

Some models (Walberg 1984 and 1986; Creemers 1994) also expect that outcome variables have an effect on the learning process and, thus, allow for a non-recursive relationship between learning process and learning outcomes. Positive or negative experiences with subject-matter learning can influence process variables such as habits and attitudes towards the learning of a subject, increase or decrease the amount of time spent on homework, and so on. Another example is long-term interest in a subject or domain, which can be the outcome of learning but also affects the students' commitment to learning.

It also needs to be recognised that vertical or horizontal relationships might not be the only explanations for differences in learning outcomes. Antecedents at the school level, for example, are often influenced by process variables at the system level like educational policies. Another example is the possibility that the socio-cultural context (antecedent at the system level) might have an influence on instructional practices (process at the classroom level), which in turn leads to differences in student outcomes.

An important corollary of the intricate relationships between the various cells in Figure 3.1 is that each one of the observed variables is likely to convey multiple information (*i.e.* both information on the dimension that the variable is intended to measure, and information on related antecedents or process variables). For example, the variables identifying the study programme or grade of the students not only contain direct information on their instructional setting and curriculum, but, in many cases, also indirect information on students' probable prior level of achievement, maybe of their home background, and possibly some of the characteristics of their teachers.

In view of the complexity of potential relationships between these variable types, explicit causal relationships were not included in this conceptual mapping. There are too many potential relationships between these components (including cross-level relationships) that might be relevant for PISA and which could not be integrated into one 'general' conceptual model.



Figure 3.2
The two-dimensional conceptual matrix with
examples of variables collected or available from other sources

| | Antecedents | Processes | Outcomes |
|--|--|--|--|
| The education system as a whole | <p>Cell 1: Macro-economic and demographic context</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Gross Domestic Product ▪ Distribution of wealth (Gini index) ▪ Percentage of immigrants | <p>Cell 5: Policies and organisation of education</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Organisation of education (school autonomy, programme structure) ▪ Teacher qualifications and training requirements ▪ School entry age, retention | <p>Cell 9: Outcomes at the level of the education system</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ System level aggregates of: reading, mathematical and scientific literacy ▪ Habits in relation to content domains ▪ Attitudinal outcomes ▪ Life skills and learning strategies ▪ Equity related outcomes |
| Educational institutions | <p>Cell 2: Characteristics of educational institutions</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ The involvement of parents ▪ Social intake ▪ Source of funding, location and size ▪ Type of educational provider (e.g. out-of-school, educational media programme) | <p>Cell 6: Institutional policies and practice</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Instructional support including both material and human resources ▪ Policies and practices, including assessment and admittance policies ▪ Activities to promote student learning | <p>Cell 10: Learning outcomes at the institutional level</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Institution level aggregates of: reading, mathematical and scientific literacy ▪ Habits in relation to content domains ▪ Affective outcomes (e.g. attitudes to mathematics) ▪ Life skills and learning strategies ▪ Differences in outcomes for students of various backgrounds |
| Instructional settings | <p>Cell 3: Characteristics of instructional settings</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Teacher qualifications ▪ Classroom size | <p>Cell 7: Learning environment</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Ability grouping ▪ Teaching styles ▪ Learning time | <p>Cell 11: Learning outcomes at the level of instructional setting</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Classroom motivation to learn ▪ Average classroom performance |
| Individual participants in education and learning | <p>Cell 4: Individual background</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Parental occupational status ▪ Parental educational level ▪ Educational resources at home ▪ Ethnicity and language ▪ Age and gender | <p>Cell 8: Individual learning process</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Engagement and attitudes to science ▪ Self-concept and self-efficacy when learning science ▪ Motivation to learn science | <p>Cell 12: Individual outcomes</p> <p><i>For example:</i></p> <ul style="list-style-type: none"> ▪ Reading, mathematical and scientific literacy ▪ Affective outcomes (e.g. attitudes to science) |



Therefore, this conceptual mapping provides a point of reference in the conceptual framework for PISA 2006 rather than as a general ‘PISA model’. More detailed models should be developed for particular research areas and for specific relationships. Relevant variables in these more specific models, however, could still be located within this conceptual two-dimensional matrix.

Figure 3.2 shows examples of variables that were collected or are available for each cell of the two-dimensional conceptual matrix that has guided the development of context questionnaire for PISA 2006.

RESEARCH AREAS IN PISA 2006

PISA’s contributions to policy makers’ and educators’ needs were maximised by identifying possible policy-relevant research areas and choosing carefully from among the many possibilities so that the strengths of the PISA design were capitalised on.

The following research areas were developed following recommendations from the Questionnaire Expert Group:

- Student’s engagement in science: In part, this research area parallels the research area on engagement in mathematics in PISA 2003. However it has been expanded to incorporate aspects of the affective dimension more comprehensively, but in a way that is not bound to a ‘cognitive unit context’. It covers self-related cognitions, motivational preferences, emotional factors as well as behaviour-related variables (such as participation in science-related activities in and out of school);
- Teaching and learning of science: This research area addresses how instructional strategies are used to teach science at school and to what extent science instruction is different across types of education and schools;
- Scientific literacy and environment: It is of interest to policy-makers to what extent schools contribute to the awareness of and attitudes toward environmental problems and challenges among 15-year-old students. This is an area related to scientific literacy (OECD, 2006) and school instruction in this area can be regarded as a potential source of information;
- Organisation of educational systems: This research area explores the relationships between scientific literacy and structural characteristics of educational systems, such as general vs. specialised curricula, comprehensive vs. tracked study programmes, centralised vs. decentralised management of schools;
- Science attainment and the labour market: The role and value of science education and scientific literacy as a preparation for future occupation are discussed in this research area, both in terms of students’ expectations and school practices concerning orientation and information for students about science-related careers.

The following two research areas had been also been developed but were not retained for the main survey after reviewing the field trial results and after the PGB decided on the priorities for the final data collection:

- Student performance and gender: This research area focused on student performance in all three major domains and comprised not only data from PISA 2006 but also from previous PISA cycles and previous international studies (IEA mathematics and science studies, IEA reading literacy studies);
- Parental investment and scientific literacy: This research area was concerned with the effects of parental involvement and parenting styles on students’ science-related career expectations and scientific literacy.

Table 3.1 shows for each research area the main constructs and variables that were included in the PISA 2006 main data collection to explore each of the research areas.



Table 3.1
Themes and constructs/variables in PISA 2006

| Research area | Constructs or variables |
|--|---|
| Student engagement in science | Science self-efficacy (StQ) Science self-concept (StQ) Interest in learning science (StQ) Enjoyment of learning science (StQ) Instrumental motivation to learn science (StQ) Future-oriented science motivation (StQ) General value of science (StQ) Students' personal value of science (StQ) Students' science-related activities (StQ) Parents' general value of science (PaQ) Parents' personal value of science (PaQ) |
| Teaching and learning of science | Interactive science teaching (StQ) Hands-on science teaching activities (StQ) Student investigation in science lessons (StQ) Science teaching with focus on applications (StQ) Time spent on learning science (StQ) |
| Scientific literacy and the environment | Students' awareness of environmental issues (StQ) Students' perception of environmental issues (StQ) Students' environmental optimism (StQ) Responsibility for sustainable development (StQ) School activities to promote environmental learning (ScQ) Parents' perception of environmental issues (PaQ) Parents' environmental optimism (PaQ) |
| Organisation of educational systems | School size, location and funding (ScQ) Grade range (ScQ) Class size (ScQ) Grade repetition at school (ScQ) Ability grouping (ScQ) Teacher-student ratio (ScQ) Computer availability at school (ScQ) School selectivity (ScQ) School responsibility for resource allocation (ScQ) School responsibility for curriculum & assessment (ScQ) School accountability policies (ScQ) Assessment practices (ScQ) Activities to promote engagement with science learning Teacher shortage (ScQ) Quality of educational resources (ScQ) Parents' perception of school quality (PaQ) |
| Science attainment and the labour market | School preparation for science career (StQ) School information on science careers (StQ) Expected occupation at 30 (StQ) Career preparation at school (ScQ) Student's science activities at age 10 (PaQ) Parents' views on importance of science (PaQ) Parents' view on student's science career motivation (PaQ) |

Note: StQ = Student questionnaire; ScQ = School questionnaire; PaQ = Parent questionnaire.



THE DEVELOPMENT OF THE CONTEXT QUESTIONNAIRES

From the theoretical bases of each research area, as elaborated, a large number of constructs were defined and their measurement operationalised through obtaining or writing questionnaire items (often in item batteries to form scales).

Small scale trials were undertaken in a range of countries and languages. Firstly a pre-pilot with a small convenience sample was undertaken in Australia. It involved a think aloud process where students were asked to complete the questionnaire while verbalising their thought processes. The pre-pilot provided qualitative feedback on the understanding and appropriateness of the items. After refining the items in light of the pre-pilot results, a series of pilot studies was undertaken in Japan (Japanese), Germany (German), Canada (French) and Australia (English). The pilots consisted of collecting questionnaire data from small convenience samples in each country. After data collection, students were collectively interviewed about their understanding of each question, particularly probing for relevance and ambiguity. The pilot therefore yielded both quantitative and qualitative data, plus conducting group interviews on the questions.

After further refinement of the questions, data was gathered in 2005 from a full scale field trial of student, school and parent questionnaires in each of the 57 participating countries in over 40 languages. The field trial was able to facilitate the investigation of a large number of student questionnaire items through the use of a rotational design with four questionnaire forms that were randomly allocated to students.

In addition, the field trial was used for in-depth analysis of the following aspects:

- Two sets of items were trialled as dichotomous and Likert-type items in parallel forms to explore cross-cultural differences in responses to either item type. Results showed some tendencies to more extreme responses in some countries but on balance it seemed more appropriate to use Likert-type items in the PISA questionnaires (Walker, 2006; Walker, 2007);
- Two sets of items were trialled with different category headings: Nine items measuring control strategies for science learning were trialled in one version asking about frequencies and in another one asking about agreement. Seven items measuring student participation in activities to protect the environment were trialled both with categories reflecting frequencies and with categories reflecting both frequency and intent. The field trial data were analysed to decide on the more appropriate but neither set of items was included in the final main study questionnaire;
- Two different sets of items measuring science self-efficacy were trialled. One set of items included asked about student confidence in tasks related to general science understanding, the other set about student confidence in doing science subject-specific tasks. Both of the item sets had good scaling characteristics and it was decided to retain the items measuring self-confidence in general science tasks due to a better fit with the science literacy framework;
- Student and parent questionnaire data were used to explore the consistency of responses regarding parental education and occupation. Results showed relatively high consistency between student and parent reports on occupation but somewhat lower consistencies for data on educational levels (Schulz, 2006).

Empirical analyses included the examination of:

- The frequency of missing values by country;
- The magnitude and consistency of item-total score correlations for each scale, by country;



- The magnitude and the consistency of scale reliability (Cronbach's Alpha), by country;
- The magnitude and consistency of correlations with each scale and science achievement as determined in the PISA field trial science test, by country;
- Exploratory and confirmatory factor analyses to determine construct validity and reliability of each scale across the pooled sample;
- Multiple-group models to assess the parameter invariance of factor models across countries;
- Item Response Theory (IRT) analyses to determine item fit for the pooled sample;
- Item-by-country interaction of items across countries using IRT scaling.

In addition to the empirical analyses, the choice of items, item format and wording was informed by:

- Direction from the PISA Governing Board;
- Feedback from National Project Managers;
- Feedback from linguistic experts;
- Discussions with the Questionnaire Expert Group;
- Discussions with members of the Science Expert Group;
- Consultation with science forum nominees of the PISA Governing Board;
- Consultation with the OECD secretariat.

Finally, in October 2005 a large and comprehensive set of potential items and topics was provided to the PISA Governing Board. From this set, the PGB indicated priority areas for investigation.

THE COVERAGE OF THE QUESTIONNAIRE MATERIAL

Student questionnaire

The student questionnaire was administered after the literacy assessment and it took students about 30 minutes to complete the instrument. The core questions on home background were similar to those used in PISA 2003, however, for some questions the wording was modified to improve the quality of the data collection based on experiences in previous surveys. Appendix 5 lists the core questions with changes in wording from PISA 2003 to PISA 2006.

The questionnaire covered the following aspects:

- Student characteristics: Grade, study programme, age and gender;
- Family background: Occupation of parents, education of parents, home possessions, number of books at home, country of birth for student and parents, language spoken at home;
- Students' views on science: Enjoyment of science, confidence in solving science tasks, general and personal value of science, participation in science-related activities, sources of information on science and general interest in learning science;
- Students; views on the environment: Awareness of environmental issues, source of information on the environment, perception of the impact of environmental issues, optimism about environmental issues and sense of responsibility for sustainable development;



- Students' views of science-related careers: Usefulness of schooling as preparation for the science labour market, information about science-related careers, future-oriented motivations for science and expected occupation at 30;
- Students' reports on learning time: Mode and duration of students' learning time in different subject areas and duration of students' out-of-school lessons;
- Students' views on teaching and learning of science: Science course taking in current and previous year, nature of science teaching at school (interactive, hands-on activities, student investigations and use of applications), future-oriented motivations to learn science, importance of doing well in subject areas (science, mathematics and test language subjects) and academic self-concept in science.

School questionnaire

The school questionnaire was administered to the school principal and took about 20 minutes to be completed. It covered a variety of school-related aspects:

- Structure and organisation of the school: Enrolment, ownership, funding, grade levels, grade repetition, average test language class size, community size and tracking/ability grouping;
- Staffing and management: Number of teachers, availability of science teaching staff, responsibility for decision-making at school and influences of external bodies on school-level decisions;
- The school's resources: Number of computers at school and principals' views on quality and quantity of staffing and educational resources;
- Accountability and admission practices: Accountability to parents, parental pressure on school, use of achievement data, parental choice of local school(s) and school admittance policies;
- Teaching of science and the environmental issues: School activities to promote learning of science, environmental issues in school curriculum and school activities to promote learning of environmental issues; and
- Aspects of career guidance: Students' opportunities to participate in career information activities, student training through local businesses, influence of business on school curriculum and structure of career guidance at school.

International options

As in previous surveys, additional questionnaire material was developed, which was offered as international options to participating countries. In PISA 2006, two international options were available, the ICT Familiarity questionnaire and the parent questionnaire.

Information communication technology (ICT) familiarity questionnaire

The ICT familiarity questionnaire consisted of questions regarding the students' use of, familiarity with and attitudes towards information communication technology which was defined as the use of any equipment or software for processing or transmitting digital information that performs diverse general functions whose options can be specified or programmed by its user. The questionnaire was administered to students after the international student questionnaire (sometimes combined within the same booklet) and it took about five minutes to be completed. It covered the following ICT-related aspects:

- Use of ICT: Students' experience with computers at different locations and frequency of ICT use for different purposes;
- Affective responses to ICT: Confidence in carrying out ICT-related tasks.



Parent questionnaire

The parent questionnaire covered both parental social background and aspects related to some of the research areas. It took about ten minutes to complete and one questionnaire was administered per student. The questionnaire covered the following aspects:

- Parental reports related to school and science learning: The students' past science activities, parental perceptions of value and quality of the student's schooling, parental views on science-related careers and parental general and personal value of science;
- Parental views on the environment: Parental awareness of environmental views and environmental optimism;
- Annual spending on children's education;
- Parental background: Age, occupation (both parents), education (both parents) and household income.

National questionnaire material

National centres could add nationally specific items to any of the questionnaires. Insertion of national items into the international questionnaires had to be agreed upon with the international study centre during the review of adaptations. National student questionnaire options, which took no longer than ten minutes to be completed, could be administered after the international student questionnaire and international options. If the length of the additional material exceeded ten minutes, national centres were requested to administer their national questionnaire material in follow-up sessions.

THE IMPLEMENTATION OF THE CONTEXT QUESTIONNAIRES

In order to make questions understood by 15-year-old students, their parents and school principals in participating countries, it was necessary to adapt parts of the questionnaire material from the international source version to the national context without jeopardising the comparability of the collected data. This is particularly important for questions that relate to specific aspects of educational systems like educational levels, study programmes or certain school characteristics which differ in terminology across countries.

To achieve maximum comparability, a process was implemented during which each adaptation was reviewed and discussed by the international study centre and national study centres. To facilitate this process, national centres were asked to complete a questionnaire adaptation spreadsheet (QAS), where adaptations to the questionnaire material were documented.

Each adaptation had to be reviewed and agreed upon before the questionnaire material could be submitted for linguistic verification and the final optical check (see Chapter 5). The QAS also contained information about additional national questionnaire material and any deviation from the international questionnaire format.

Prior to the review of questionnaire adaptations, national centres were asked to complete three different tables describing necessary adaptations:

- Study programme tables (STP): These document the range of different study programmes that are available for 15-year-old students across participating countries. This information was not only used as a codebook to collect these data from school records but also assisted the review of questionnaire adaptations;
- Language tables (LNT): These document the language categories included in the question about language use at home; and
- Country tables (CNT): These document the country categories in the questions about the country of birth for students and parents.



Information on parental occupation and the students' expected occupation was collected through open-ended questions both in student and parent questionnaires. The responses were then coded according to the International Standard Classification of Occupations (ISCO) (International Labour Organisation, 1990). Once occupations had been coded into ISCO, the codes were re-coded into the International Socio-Economic Index of Occupational Status (*ISEI*) (Ganzeboom, de Graaf & Treiman, 1992), which provides a measure of the socio-economic status of occupations comparable across the countries participating in PISA.

The International Standard Classification of Education (ISCED) (OECD, 1999) was used as a typology to classify educational qualifications and study programmes. The ISCED classification was used to get comparable data across countries. Whereas this information was readily available for OECD member countries, for partner countries and economies extensive reviews of their educational systems in cooperation with national centres were necessary to map educational levels to the ISCED framework.



Reader's Guide

Country codes – the following country codes are used in this report:

OECD countries

| | |
|-----|---------------------------------|
| AUS | Australia |
| AUT | Austria |
| BEL | Belgium |
| BEF | Belgium (French Community) |
| BEN | Belgium (Flemish Community) |
| CAN | Canada |
| CAE | Canada (English Community) |
| CAF | Canada (French Community) |
| CZE | Czech Republic |
| DNK | Denmark |
| FIN | Finland |
| FRA | France |
| DEU | Germany |
| GRC | Greece |
| HUN | Hungary |
| ISL | Iceland |
| IRL | Ireland |
| ITA | Italy |
| JPN | Japan |
| KOR | Korea |
| LUX | Luxembourg |
| LXF | Luxembourg (French Community) |
| LXG | Luxembourg (German Community) |
| MEX | Mexico |
| NLD | Netherlands |
| NZL | New Zealand |
| NOR | Norway |
| POL | Poland |
| PRT | Portugal |
| SVK | Slovak Republic |
| ESP | Spain |
| ESB | Spain (Basque Community) |
| ESC | Spain (Catalonian Community) |
| ESS | Spain (Castillian Community) |
| SWE | Sweden |
| CHE | Switzerland |
| CHF | Switzerland (French Community) |
| CHG | Switzerland (German Community) |
| CHI | Switzerland (Italian Community) |

| | |
|-----|----------------|
| TUR | Turkey |
| GBR | United Kingdom |
| IRL | Ireland |
| SCO | Scotland |
| USA | United States |

Partner countries and economies

| | |
|-----|----------------------------|
| ARG | Argentina |
| AZE | Azerbaijan |
| BGR | Bulgaria |
| BRA | Brazil |
| CHL | Chile |
| COL | Colombia |
| EST | Estonia |
| HKG | Hong Kong-China |
| HRV | Croatia |
| IDN | Indonesia |
| JOR | Jordan |
| KGZ | Kyrgyzstan |
| LIE | Liechtenstein |
| LTU | Lithuania |
| LVA | Latvia |
| LVL | Latvia (Latvian Community) |
| LVR | Latvia (Russian Community) |
| MAC | Macao-China |
| MNE | Montenegro |
| QAT | Qatar |
| ROU | Romania |
| RUS | Russian Federation |
| SRB | Serbia |
| SVN | Slovenia |
| TAP | Chinese Taipei |
| THA | Thailand |
| TUN | Tunisia |
| URY | Uruguay |



References

- Adams, R.J., Wilson, M. & Wang, W.C.** (1997), The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, No. 21, pp. 1-23.
- Adams, R.J., Wilson, M. R. & Wu, M.L.** (1997), Multilevel item response models: An approach to errors in variables regression, *Journal of Educational and Behavioural Statistics*, No. 22 (1), pp. 46-75.
- Adams, R.J. & Wu, M.L.** (2002), *PISA 2000 Technical Report*, OECD, Paris.
- Bollen, K.A. & Long, S.J.** (1993) (eds.), *Testing Structural Equation Models*, Newbury Park: London.
- Beaton, A.E.** (1987), Implementing the new design: The NAEP 1983-84 technical report (Rep. No. 15-TR-20), Princeton, NJ: Educational Testing Service.
- Buchmann, C.** (2000), Family structure, parental perceptions and child labor in Kenya: What factors determine who is enrolled in school? *Soc. Forces*, No. 78, pp. 1349-79.
- Buchmann, C.** (2002), Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges, in Porter, A.C. and Gamoran, A. (eds.). *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 150-97), Washington, DC: National Academy Press.
- Creemers, B.P.M.** (1994), *The Effective Classroom*, London: Cassell.
- Cochran, W.G.** (1977), *Sampling techniques*, third edition, New York, NY: John Wiley and Sons.
- Ganzeboom, H.B.G., de Graaf, P.M. & Treiman, D.J.** (1992), A standard international socio-economic index of occupational status, *Social Science Research*, No. 21, pp. 1-56.
- Ganzeboom H.B. & Treiman, D.J.** (1996), Internationally comparable measures of occupational status for the 1988 international standard classification of occupations, *Social Science Research*, No. 25, pp. 201-239.
- Grisay, A.** (2003), Translation procedures in OECD/PISA 2000 international assessment, *Language Testing*, No. 20 (2), pp. 225-240.
- Hambleton, R.K., Swaminathan, H. & Rogers, H.J.** (1991), *Fundamentals of item response theory*, Newbury Park, London, New Delhi: SAGE Publications.
- Hambleton, R.K., Merenda, P.F. & Spielberger, C.D.** (2005), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*, IEA Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.
- Harkness, J.A., Van de Vijver, F.J.R. & Mohler, P.Ph** (2003), *Cross-Cultural Survey Methods*, Wiley-Interscience, John Wiley & Sons, Inc., Hoboken, New Jersey.
- Harvey-Beavis, A.** (2002), Student and School Questionnaire Development, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report*, (pp. 33-38), OECD, Paris.
- International Labour Organisation (ILO)** (1990), *International Standard Classification of Occupations: ISCO-88*. Geneva: International Labour Office.
- Jöreskog, K.G. & Sörbom, Dag** (1993), *LISREL 8 User's Reference Guide*, Chicago: SSI.
- Judkins, D.R.** (1990), Fay's Method of Variance Estimation, *Journal of Official Statistics*, No. 6 (3), pp. 223-239.
- Kaplan, D.** (2000), *Structural equation modeling: Foundation and extensions*, Thousand Oaks: SAGE Publications.
- Keyfitz, N.** (1951), Sampling with probabilities proportionate to science: Adjustment for changes in probabilities, *Journal of the American Statistical Association*, No. 46, American Statistical Association, Alexandria, pp. 105-109.
- Kish, L.** (1992), Weighting for Unequal, *Pi. Journal of Official Statistics*, No. 8 (2), pp. 183-200.
- LISREL** (1993), K.G. Jöreskog & D. Sörbom, [computer software], Lincolnwood, IL: Scientific Software International, Inc.
- Lohr, S.L.** (1999), *Sampling: Design and Analysis*, Duxberry: Pacific Grove.
- Macaskill, G., Adams, R.J. & Wu, M.L.** (1998), Scaling methodology and procedures for the mathematics and science literacy, advanced mathematics and physics scale, in M. Martin and D.L. Kelly, Editors, *Third International Mathematics and Science Study, technical report Volume 3: Implementation and analysis*, Boston College, Chestnut Hill, MA.
- Masters, G.N. & Wright, B.D.** (1997), The Partial Credit Model, in W.J. van der Linden, & R.K. Hambleton (eds.), *Handbook of Modern Item Response Theory* (pp. 101-122), New York/Berlin/Heidelberg: Springer.

- Mislevy, R.J.** (1991), Randomization-based inference about latent variables from complex samples, *Psychometrika*, No. 56, pp. 177-196.
- Mislevy, R.J., Beaton, A., Kaplan, B.A. & Sheehan, K.** (1992), Estimating population characteristics from sparse matrix samples of item responses, *Journal of Educational Measurement*, No. 29 (2), pp. 133-161.
- Mislevy, R.J. & Sheehan, K.M.** (1987), Marginal estimation procedures, in Beaton, A.E., Editor, 1987. *The NAEP 1983-84 technical report*, National Assessment of Educational Progress, Educational Testing Service, Princeton, pp. 293-360.
- Mislevy, R.J. & Sheehan, K.M.** (1989), Information matrices in latent-variable models, *Journal of Educational Statistics*, No. 14, pp. 335-350.
- Mislevy, R.J. & Sheehan, K.M.** (1989), The role of collateral information about examinees in item parameter estimation, *Psychometrika*, No. 54, pp. 661-679.
- Monseur, C. & Berezner, A.** (2007), The Computation of Equating Errors in International Surveys in Education, *Journal of Applied Measurement*, No. 8 (3), 2007, pp. 323-335.
- Monseur, C.** (2005), An exploratory alternative approach for student non response weight adjustment, *Studies in Educational Evaluation*, No. 31 (2-3), pp. 129-144.
- Muthen, B. & L. Muthen** (1998), [computer software], *Mplus* Los Angeles, CA: Muthen & Muthen.
- Muthen, B., du Toit, S.H.C. & Spisic, D.** (1997), *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes*, unpublished manuscript.
- OECD** (1999), *Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries*, OECD, Paris.
- OECD** (2003), *Literacy Skills for the World of Tomorrow: Further results from PISA 2000*, OECD, Paris.
- OECD** (2004), *Learning for Tomorrow's World – First Results from PISA 2003*, OECD, Paris.
- OECD** (2005), *Technical Report for the OECD Programme for International Student Assessment 2003*, OECD, Paris.
- OECD** (2006), *Assessing Scientific, Reading and Mathematical Literacy: A framework for PISA 2006*, OECD, Paris.
- OECD** (2007), *PISA 2006: Science Competencies for Tomorrow's World*, OECD, Paris.
- PISA Consortium** (2006), *PISA 2006 Main Study Data Management Manual*, https://mypisa.acer.edu.au/images/mypisadoc/opmanual/pisa2006_data_management_manual.pdf
- Rasch, G.** (1960), Probabilistic models for some intelligence and attainment tests, Copenhagen: Nielsen & Lydiche.
- Routitski A. & Berezner, A.** (2006), Issues influencing the validity of cross-national comparisons of student performance. Data Entry Quality and Parameter Estimation. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in San Francisco, 7-11 April, https://mypisa.acer.edu.au/images/mypisadoc/era06routitsky_berezner.pdf
- Rust, K.** (1985), Variance Estimation for Complex Estimators in Sample Surveys, *Journal of Official Statistics*, No. 1, pp. 381-397.
- Rust, K.F. & Rao, J.N.K.** (1996), Variance Estimation for Complex Surveys Using Replication Techniques, *Survey Methods in Medical Research*, No. 5, pp. 283-310.
- Shao, J.** (1996), Resampling Methods in Sample Surveys (with Discussion), *Statistics*, No. 27, pp. 203-254.
- Särndal, C.-E., Swensson, B. & Wretman, J.** (1992), *Model Assisted Survey Sampling*, New York: Springer-Verlag.
- SAS® CALIS** (1992), W. Hartmann [computer software], Cary, NC: SAS Institute Inc.
- Scheerens, J.** (1990), School effectiveness and the development of process indicators of school functioning, *School effectiveness and school improvement*, No. 1, pp. 61-80.
- Scheerens, J. & Bosker, R.J.** (1997), *The Foundations of School Effectiveness*, Oxford: Pergamon.
- Schulz, W.** (2002), Constructing and Validating the Questionnaire composites, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report*, OECD, Paris.
- Schulz, W.** (2004), Mapping Student Scores to Item Responses, in W. Schulz and H. Sibberns (eds.), *IEA Civic Education Study, Technical Report* (pp. 127-132), Amsterdam: IEA.
- Schulz, W.** (2006a), *Testing Parameter Invariance for Questionnaire Indices using Confirmatory Factor Analysis and Item Response Theory*, Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.
- Schulz, W.** (2006b), *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003*, Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.
- Thorndike, R.L.** (1973), *Reading comprehension in fifteen countries*, New York, Wiley: and Stockholm: Almqvist & Wiksell.
- Travers, K.J. & Westbury, I.** (1989), *The IEA Study of Mathematics I: Analysis of Mathematics Curricula*, Oxford: Pergamon Press.



- Travers, K.J., Garden R.A. & Rosier, M.** (1989), Introduction to the Study, in Robitaille, D. A. and Garden, R. A. (eds), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics Curricula*, Oxford: Pergamon Press.
- Verhelst, N.** (2002), Coder and Marker Reliability Studies, in R.J. Adams & M.L. Wu (eds.), *PISA 2000 Technical Report*. OECD, Paris.
- Walberg, H.J.** (1984), Improving the productivity of American schools, *Educational Leadership*, No. 41, pp. 19-27.
- Walberg, H.** (1986), Synthesis of research on teaching, in M. Wittrock (ed.), *Handbook of research on teaching* (pp. 214-229), New York: Macmillan.
- Walker, M.** (2006), *The choice of Likert or dichotomous items to measure attitudes across culturally distinct countries in international comparative educational research*. Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.
- Walker, M.** (2007), Ameliorating Culturally-Based Extreme Response Tendencies To Attitude items, *Journal of Applied Measurement*, No. 8, pp. 267-278.
- Warm, T.A.** (1989), Weighted Likelihood Estimation of Ability in Item Response Theory, *Psychometrika*, No. 54 (3), pp. 427-450.
- Westat** (2007), *WesVar[®] 5.1* Computer software and manual, Rockville, MD: Author (also see <http://www.westat.com/wesvar>).
- Wilson, M.** (1994), Comparing Attitude Across Different Cultures: Two Quantitative Approaches to Construct Validity, in M. Wilson (ed.), *Objective measurement II: Theory into practice* (pp. 271-292), Norwood, NJ: Ablex.
- Wolter, K.M.** (2007), *Introduction to Variance Estimation*. Second edition, Springer: New York.
- Wu, M.L., Adams, R.J. & Wilson, M.R.** (1997), *ConQuest[®]: Multi-Aspect Test Software* [computer program manual], Camberwell, Vic.: Australian Council for Educational Research.



List of abbreviations – the following abbreviations are used in this report:

| | | | |
|-------|---|---------|--|
| ACER | Australian Council for Educational Research | NPM | National Project Manager |
| AGFI | Adjusted Goodness-of-Fit Index | OECD | Organisation for Economic Cooperation and Development |
| BRR | Balanced Repeated Replication | PISA | Programme for International Student Assessment |
| CBAS | Computer Based Assessment of Science | PPS | Probability Proportional to Size |
| CFA | Confirmatory Factor Analysis | PGB | PISA Governing Board |
| CFI | Comparative Fit Index | PQM | PISA Quality Monitor |
| CITO | National Institute for Educational Measurement, The Netherlands | PSU | Primary Sampling Units |
| CIVED | Civic Education Study | QAS | Questionnaire Adaptations Spreadsheet |
| DIF | Differential Item Functioning | RMSEA | Root Mean Square Error of Approximation |
| ENR | Enrolment of 15-year-olds | RN | Random Number |
| ESCS | PISA Index of Economic, Social and Cultural Status | SC | School Co-ordinator |
| ETS | Educational Testing Service | SE | Standard Error |
| IAEP | International Assessment of Educational Progress | SD | Standard Deviation |
| I | Sampling Interval | SEM | Structural Equation Modelling |
| ICR | Inter-Country Coder Reliability Study | SMEG | Subject Matter Expert Group |
| ICT | Information Communication Technology | SPT | Study Programme Table |
| IEA | International Association for the Evaluation of Educational Achievement | TA | Test Administrator |
| INES | OECD Indicators of Education Systems | TAG | Technical Advisory Group |
| IRT | Item Response Theory | TCS | Target Cluster Size |
| ISCED | International Standard Classification of Education | TIMSS | Third International Mathematics and Science Study |
| ISCO | International Standard Classification of Occupations | TIMSS-R | Third International Mathematics and Science Study – Repeat |
| ISEI | International Socio-Economic Index | VENR | Enrolment for very small schools |
| MENR | Enrolment for moderately small school | WLE | Weighted Likelihood Estimates |
| MOS | Measure of size | | |
| NCQM | National Centre Quality Monitor | | |
| NDP | National Desired Population | | |
| NEP | National Enrolled Population | | |
| NFI | Normed Fit Index | | |
| NIER | National Institute for Educational Research, Japan | | |
| NNFI | Non-Normed Fit Index | | |



Table of contents

| | |
|--|-----------|
| FOREWORD | 3 |
| CHAPTER 1 PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT: AN OVERVIEW | 19 |
| Participation | 21 |
| Features of PISA | 22 |
| Managing and implementing PISA | 23 |
| Organisation of this report | 23 |
| READER'S GUIDE | 25 |
| CHAPTER 2 TEST DESIGN AND TEST DEVELOPMENT | 27 |
| Test scope and format | 28 |
| Test design | 28 |
| Test development centres | 29 |
| Development timeline | 30 |
| The PISA 2006 scientific literacy framework | 30 |
| Test development – cognitive items | 31 |
| ▪ Item development process..... | 31 |
| ▪ National item submissions..... | 33 |
| ▪ National review of items..... | 34 |
| ▪ International item review..... | 35 |
| ▪ Preparation of dual (English and French) source versions..... | 35 |
| Test development – attitudinal items | 35 |
| Field trial | 38 |
| ▪ Field trial selection..... | 38 |
| ▪ Field trial design..... | 39 |
| ▪ Despatch of field trial instruments..... | 40 |
| ▪ Field trial coder training..... | 40 |
| ▪ Field trial coder queries..... | 40 |
| ▪ Field trial outcomes..... | 41 |
| ▪ National review of field trial items..... | 42 |
| Main study | 42 |
| ▪ Main study science items..... | 42 |
| ▪ Main study reading items..... | 44 |
| ▪ Main study mathematics items..... | 45 |
| ▪ Despatch of main study instruments..... | 46 |
| ▪ Main study coder training..... | 46 |
| ▪ Main study coder query service..... | 46 |
| ▪ Review of main study item analyses..... | 47 |



| | |
|---|---------------|
| CHAPTER 3 THE DEVELOPMENT OF THE PISA CONTEXT QUESTIONNAIRES | 49 |
| Overview | 50 |
| The conceptual structure | 51 |
| ▪ A conceptual framework for PISA 2006 | 51 |
| Research areas in PISA 2006 | 55 |
| The development of the context questionnaires | 57 |
| The coverage of the questionnaire material | 58 |
| ▪ Student questionnaire | 58 |
| ▪ School questionnaire | 59 |
| ▪ International options | 59 |
| ▪ National questionnaire material | 60 |
| The implementation of the context questionnaires | 60 |
| CHAPTER 4 SAMPLE DESIGN | 63 |
| Target population and overview of the sampling design | 64 |
| Population coverage, and school and student participation rate standards | 65 |
| ▪ Coverage of the PISA international target population | 65 |
| ▪ Accuracy and precision | 66 |
| ▪ School response rates | 66 |
| ▪ Student response rates | 68 |
| Main study school sample | 68 |
| ▪ Definition of the national target population | 68 |
| ▪ The sampling frame | 69 |
| ▪ Stratification | 70 |
| ▪ Assigning a measure of size to each school | 74 |
| ▪ School sample selection | 74 |
| ▪ PISA and TIMSS or PIRLS overlap control | 76 |
| ▪ Student samples | 82 |
| CHAPTER 5 TRANSLATION AND CULTURAL APPROPRIATENESS OF THE TEST AND SURVEY MATERIAL | 85 |
| Introduction | 86 |
| Development of source versions | 86 |
| Double translation from two source languages | 87 |
| PISA translation and adaptation guidelines | 88 |
| Translation training session | 89 |
| Testing languages and translation/adaptation procedures | 89 |
| International verification of the national versions | 91 |
| ▪ VegaSuite | 93 |
| ▪ Documentation | 93 |
| ▪ Verification of test units | 93 |
| ▪ Verification of the booklet shell | 94 |
| ▪ Final optical check | 94 |
| ▪ Verification of questionnaires and manuals | 94 |
| ▪ Final check of coding guides | 95 |
| ▪ Verification outcomes | 95 |



| | |
|---|------------|
| Translation and verification outcomes – national version quality | 96 |
| ▪ Analyses at the country level..... | 96 |
| ▪ Analyses at the item level..... | 103 |
| ▪ Summary of items lost at the national level, due to translation, printing or layout errors..... | 104 |
| CHAPTER 6 FIELD OPERATIONS | 105 |
| Overview of roles and responsibilities | 106 |
| ▪ National project managers..... | 106 |
| ▪ School coordinators..... | 107 |
| ▪ Test administrators..... | 107 |
| ▪ School associates..... | 108 |
| The selection of the school sample | 108 |
| Preparation of test booklets, questionnaires and manuals | 108 |
| The selection of the student sample | 109 |
| Packaging and shipping materials | 110 |
| Receipt of materials at the national centre after testing | 110 |
| Coding of the tests and questionnaires | 111 |
| ▪ Preparing for coding..... | 111 |
| ▪ Logistics prior to coding..... | 113 |
| ▪ Single coding design..... | 115 |
| ▪ Multiple coding..... | 117 |
| ▪ Managing the process coding..... | 118 |
| ▪ Cross-national coding..... | 120 |
| ▪ Questionnaire coding..... | 120 |
| Data entry, data checking and file submission | 120 |
| ▪ Data entry..... | 120 |
| ▪ Data checking..... | 120 |
| ▪ Data submission..... | 121 |
| ▪ After data were submitted..... | 121 |
| The main study review | 121 |
| CHAPTER 7 QUALITY ASSURANCE | 123 |
| PISA quality control | 124 |
| ▪ Comprehensive operational manuals..... | 124 |
| ▪ National level implementation planning document..... | 124 |
| PISA quality monitoring | 124 |
| ▪ Field trial and main study review..... | 124 |
| ▪ Final optical check..... | 126 |
| ▪ National centre quality monitor (NCQM) visits..... | 126 |
| ▪ PISA quality monitor (PQM) visits..... | 126 |
| ▪ Test administration..... | 127 |
| ▪ Delivery..... | 128 |
| CHAPTER 8 SURVEY WEIGHTING AND THE CALCULATION OF SAMPLING VARIANCE | 129 |
| Survey weighting | 130 |
| The school base weight | 131 |
| ▪ The school weight trimming factor..... | 132 |

| | |
|---|------------|
| <ul style="list-style-type: none"> ▪ The student base weight 132 ▪ School non-response adjustment..... 132 ▪ Grade non-response adjustment..... 134 ▪ Student non-response adjustment..... 135 ▪ Trimming student weights..... 136 ▪ Comparing the PISA 2006 student non-response adjustment strategy with the strategy used for PISA 2003 136 ▪ The comparison..... 138 | |
| Calculating sampling variance | 139 |
| <ul style="list-style-type: none"> ▪ The balanced repeated replication variance estimator..... 139 ▪ Reflecting weighting adjustments..... 141 ▪ Formation of variance strata..... 141 ▪ Countries where all students were selected for PISA..... 141 | |
| CHAPTER 9 SCALING PISA COGNITIVE DATA | 143 |
| The mixed coefficients multinomial logit model | 144 |
| <ul style="list-style-type: none"> ▪ The population model..... 145 ▪ Combined model..... 146 | |
| Application to PISA | 146 |
| <ul style="list-style-type: none"> ▪ National calibrations..... 146 ▪ National reports..... 147 ▪ International calibration 153 ▪ Student score generation..... 153 | |
| Booklet effects | 155 |
| Analysis of data with plausible values | 156 |
| Developing common scales for the purposes of trends | 157 |
| <ul style="list-style-type: none"> ▪ Linking PISA 2003 and PISA 2006 for reading and mathematics 158 ▪ Uncertainty in the link..... 158 | |
| CHAPTER 10 DATA MANAGEMENT PROCEDURES | 163 |
| Introduction | 164 |
| KeyQuest | 167 |
| Data management at the national centre | 167 |
| <ul style="list-style-type: none"> ▪ National modifications to the database 167 ▪ Student sampling with <i>KeyQuest</i>..... 167 ▪ Data entry quality control 167 | |
| Data cleaning at ACER | 171 |
| <ul style="list-style-type: none"> ▪ Recoding of national adaptations..... 171 ▪ Data cleaning organisation..... 171 ▪ Cleaning reports..... 171 ▪ General recodings..... 171 | |
| Final review of the data | 172 |
| <ul style="list-style-type: none"> ▪ Review of the test and questionnaire data 172 ▪ Review of the sampling data 172 | |
| Next steps in preparing the international database | 172 |



| | |
|---|------------|
| CHAPTER 11 SAMPLING OUTCOMES | 175 |
| Design effects and effective sample sizes | 187 |
| ▪ Variability of the design effect..... | 191 |
| ▪ Design effects in PISA for performance variables..... | 191 |
| Summary analyses of the design effect | 203 |
| ▪ Countries with outlying standard errors..... | 205 |
| | |
| CHAPTER 12 SCALING OUTCOMES | 207 |
| International characteristics of the item pool | 208 |
| ▪ Test targeting..... | 208 |
| ▪ Test reliability..... | 208 |
| ▪ Domain inter-correlations..... | 208 |
| ▪ Science scales..... | 215 |
| Scaling outcomes | 216 |
| ▪ National item deletions..... | 216 |
| ▪ International scaling..... | 219 |
| ▪ Generating student scale scores..... | 219 |
| Test length analysis | 219 |
| Booklet effects | 221 |
| ▪ Overview of the PISA cognitive reporting scales..... | 232 |
| ▪ PISA overall literacy scales..... | 234 |
| ▪ PISA literacy scales..... | 234 |
| ▪ Special purpose scales..... | 234 |
| Observations concerning the construction of the PISA overall literacy scales | 235 |
| ▪ Framework development..... | 235 |
| ▪ Testing time and item characteristics..... | 236 |
| ▪ Characteristics of each of the links..... | 237 |
| Transforming the plausible values to PISA scales | 246 |
| ▪ Reading..... | 246 |
| ▪ Mathematics..... | 246 |
| ▪ Science..... | 246 |
| ▪ Attitudinal scales..... | 247 |
| Link error | 247 |
| | |
| CHAPTER 13 CODING AND MARKER RELIABILITY STUDIES | 249 |
| Homogeneity analyses | 251 |
| Multiple marking study outcomes (variance components) | 254 |
| ▪ Generalisability coefficients..... | 254 |
| International coding review | 261 |
| ▪ Background to changed procedures for PISA 2006..... | 261 |
| ▪ ICR procedures..... | 261 |
| ▪ Outcomes..... | 264 |
| ▪ Cautions..... | 270 |



| | |
|---|------------|
| CHAPTER 14 DATA ADJUDICATION | 271 |
| Introduction | 272 |
| ▪ Implementing the standards – quality assurance | 272 |
| ▪ Information available for adjudication | 273 |
| ▪ Data adjudication process | 273 |
| General outcomes | 274 |
| ▪ Overview of response rate issues | 274 |
| ▪ Detailed country comments | 275 |
| | |
| CHAPTER 15 PROFICIENCY SCALE CONSTRUCTION | 283 |
| Introduction | 284 |
| Development of the described scales | 285 |
| ▪ Stage 1: Identifying possible scales | 285 |
| ▪ Stage 2: Assigning items to scales | 286 |
| ▪ Stage 3: Skills audit | 286 |
| ▪ Stage 4: Analysing field trial data | 286 |
| ▪ Stage 5: Defining the dimensions | 287 |
| ▪ Stage 6: Revising and refining with main study data | 287 |
| ▪ Stage 7: Validating | 287 |
| Defining proficiency levels | 287 |
| Reporting the results for PISA science | 290 |
| ▪ Building an item map | 290 |
| ▪ Levels of scientific literacy | 292 |
| ▪ Interpreting the scientific literacy levels | 299 |
| | |
| CHAPTER 16 SCALING PROCEDURES AND CONSTRUCT VALIDATION OF CONTEXT QUESTIONNAIRE DATA | 303 |
| Overview | 304 |
| Simple questionnaire indices | 304 |
| ▪ Student questionnaire indices | 304 |
| ▪ School questionnaire indices | 307 |
| ▪ Parent questionnaire indices | 309 |
| Scaling methodology and construct validation | 310 |
| ▪ Scaling procedures | 310 |
| ▪ Construct validation | 312 |
| ▪ Describing questionnaire scale indices | 314 |
| Questionnaire scale indices | 315 |
| ▪ Student scale indices | 315 |
| ▪ School questionnaire scale indices | 340 |
| ▪ Parent questionnaire scale indices | 342 |
| ▪ The PISA index of economic, social and cultural status (ESCS) | 346 |
| | |
| CHAPTER 17 VALIDATION OF THE EMBEDDED ATTITUDINAL SCALES | 351 |
| Introduction | 352 |
| International scalability | 353 |
| ▪ Analysis of item dimensionality with exploratory and confirmatory factor analysis | 353 |
| ▪ Fit to item response model | 353 |



| | |
|---|------------|
| ▪ Reliability..... | 355 |
| ▪ Differential item functioning..... | 355 |
| ▪ Summary of scalability..... | 357 |
| Relationship and comparisons with other variables..... | 357 |
| ▪ Within-country student level correlations with achievement and selected background variables..... | 358 |
| ▪ Relationships between embedded scales and questionnaire..... | 360 |
| ▪ Country level correlations with achievement and selected background variables..... | 361 |
| ▪ Variance decomposition..... | 363 |
| ▪ Observations from other cross-national data collections..... | 363 |
| ▪ Summary of relations with other variables..... | 364 |
| Conclusion..... | 364 |
| CHAPTER 18 INTERNATIONAL DATABASE..... | 367 |
| Files in the database..... | 368 |
| ▪ Student files..... | 368 |
| ▪ School file..... | 370 |
| ▪ Parent file..... | 370 |
| Records in the database..... | 371 |
| ▪ Records included in the database..... | 371 |
| ▪ Records excluded from the database..... | 371 |
| Representing missing data..... | 371 |
| How are students and schools identified?..... | 372 |
| Further information..... | 373 |
| REFERENCES..... | 375 |
| APPENDICES..... | 379 |
| Appendix 1 PISA 2006 main study item pool characteristics..... | 380 |
| Appendix 2 Contrast coding used in conditioning..... | 389 |
| Appendix 3 Design effect tables..... | 399 |
| Appendix 4 Changes to core questionnaire items from 2003 to 2006..... | 405 |
| Appendix 5 Mapping of ISCED to years..... | 411 |
| Appendix 6 National household possession items..... | 412 |
| Appendix 7 Exploratory and confirmatory factor analyses for the embedded items..... | 414 |
| Appendix 8 PISA consortium, staff and consultants..... | 416 |



LIST OF BOXES

| | | |
|---------|---------------------------------|----|
| Box 1.1 | Core features of PISA 2006..... | 22 |
|---------|---------------------------------|----|

LIST OF FIGURES

| | | |
|--------------|---|-----|
| Figure 2.1 | Main study Interest in Science item..... | 36 |
| Figure 2.2 | Main study Support for Scientific Enquiry item..... | 36 |
| Figure 2.3 | Field trial Match-the-opinion Responsibility item..... | 37 |
| Figure 3.1 | Conceptual grid of variable types..... | 52 |
| Figure 3.2 | The two-dimensional conceptual matrix with examples of variables collected or available from other sources..... | 54 |
| Figure 4.1 | School response rate standard..... | 67 |
| Figure 6.1 | Design for the single coding of science and mathematics..... | 115 |
| Figure 6.2 | Design for the single coding of reading..... | 116 |
| Figure 9.1 | Example of item statistics in Report 1..... | 148 |
| Figure 9.2 | Example of item statistics in Report 2..... | 149 |
| Figure 9.3 | Example of item statistics shown in Graph B..... | 150 |
| Figure 9.4 | Example of item statistics shown in Graph C..... | 151 |
| Figure 9.5 | Example of item statistics shown in Table D..... | 151 |
| Figure 9.6 | Example of summary of dodgy items for a country in Report 3a..... | 152 |
| Figure 9.7 | Example of summary of dodgy items in Report 3b..... | 152 |
| Figure 10.1 | Data management in relation to other parts of PISA..... | 164 |
| Figure 10.2 | Major data management stages in PISA..... | 166 |
| Figure 10.3 | Validity reports - general hierarchy..... | 170 |
| Figure 11.1 | Standard error on a mean estimate depending on the intraclass correlation..... | 188 |
| Figure 11.2 | Relationship between the standard error for the science performance mean and the intraclass correlation within explicit strata (PISA 2006)..... | 205 |
| Figure 12.1 | Item plot for mathematics items..... | 210 |
| Figure 12.2 | Item plot for reading items..... | 211 |
| Figure 12.3 | Item plot for science items..... | 212 |
| Figure 12.4 | Item plot for interest items..... | 213 |
| Figure 12.5 | Item plot for support items..... | 214 |
| Figure 12.6 | Scatter plot of per cent correct for reading link items in PISA 2000 and PISA 2003..... | 238 |
| Figure 12.7 | Scatter plot of per cent correct for reading link items in PISA 2003 and PISA 2006..... | 240 |
| Figure 12.8 | Scatter plot of per cent correct for mathematics link items in PISA 2003 and PISA 2006..... | 242 |
| Figure 12.9 | Scatter plot of per cent correct for science link items in PISA 2000 and PISA 2003..... | 244 |
| Figure 12.10 | Scatter plot of per cent correct for science link items in PISA 2003 and PISA 2006..... | 245 |



| | | |
|-------------|--|-----|
| Figure 13.1 | Variability of the homogeneity indices for science items in field trial | 250 |
| Figure 13.2 | Average of the homogeneity indices for science items in field trial and main study | 251 |
| Figure 13.3 | Variability of the homogeneity indices for each science item in the main study | 252 |
| Figure 13.4 | Variability of the homogeneity indices for each reading item in the main study | 252 |
| Figure 13.5 | Variability of the homogeneity indices for each mathematics item | 252 |
| Figure 13.6 | Variability of the homogeneity indices for the participating countries in the main study | 253 |
| Figure 13.7 | Example of ICR report (reading) | 269 |
| <hr/> | | |
| Figure 14.1 | Attained school response rates | 274 |
| <hr/> | | |
| Figure 15.1 | The relationship between items and students on a proficiency scale | 285 |
| Figure 15.2 | What it means to be at a level | 289 |
| Figure 15.3 | A map for selected science items | 291 |
| Figure 15.4 | Summary descriptions of the six proficiency levels on the science scale | 294 |
| Figure 15.5 | Summary descriptions of six proficiency levels in <i>identifying scientific issues</i> | 295 |
| Figure 15.6 | Summary descriptions of six proficiency levels in <i>explaining phenomena scientifically</i> | 297 |
| Figure 15.7 | Summary descriptions of six proficiency levels in <i>using scientific evidence</i> | 300 |
| <hr/> | | |
| Figure 16.1 | Summed category probabilities for fictitious item | 314 |
| Figure 16.2 | Fictitious example of an item map | 315 |
| Figure 16.3 | Scatterplot of country means for ESCS 2003 and ESCS 2006 | 347 |
| <hr/> | | |
| Figure 17.1 | Distribution of item fit mean square statistics for embedded attitude items | 354 |
| Figure 17.2 | An example of the ESC plot for item S408RNA | 356 |
| Figure 17.3 | Scatterplot of mean mathematics interest against mean mathematics for PISA 2003 | 363 |

LIST OF TABLES

| | | |
|------------|---|----|
| Table 1.1 | PISA 2006 participants | 21 |
| <hr/> | | |
| Table 2.1 | Cluster rotation design used to form test booklets for PISA 2006 | 29 |
| Table 2.2 | Test development timeline for PISA 2006 | 30 |
| Table 2.3 | Science field trial all items | 39 |
| Table 2.4 | Allocation of item clusters to test booklets for field trial | 39 |
| Table 2.5 | Science main study items (item format by competency) | 43 |
| Table 2.6 | Science main study items (item format by knowledge type) | 44 |
| Table 2.7 | Science main study items (knowledge category by competency) | 44 |
| Table 2.8 | Reading main study items (item format by aspect) | 44 |
| Table 2.9 | Reading main study items (item format by text format) | 45 |
| Table 2.10 | Reading main study items (text type by aspect) | 45 |
| Table 2.11 | Mathematics main study items (item format by competency cluster) | 45 |
| Table 2.12 | Mathematics main study items (item format by content category) | 46 |
| Table 2.13 | Mathematics main study items (content category by competency cluster) | 46 |

| | | |
|-------------|--|-----|
| Table 3.1 | Themes and constructs/variables in PISA 2006..... | 56 |
| Table 4.1 | Stratification variables | 71 |
| Table 4.2 | Schedule of school sampling activities | 78 |
| Table 5.1 | Countries sharing a common version with national adaptations | 90 |
| Table 5.2 | PISA 2006 translation/adaptation procedures..... | 91 |
| Table 5.3 | Mean deviation and root mean squared error of the item by country interactions for each version..... | 97 |
| Table 5.4 | Correlation between national item parameter estimates for Arabic versions..... | 99 |
| Table 5.5 | Correlation between national item parameter estimates for Chinese versions..... | 99 |
| Table 5.6 | Correlation between national item parameter estimates for Dutch versions..... | 99 |
| Table 5.7 | Correlation between national item parameter estimates for English versions..... | 99 |
| Table 5.8 | Correlation between national item parameter estimates for French versions..... | 99 |
| Table 5.9 | Correlation between national item parameter estimates for German versions..... | 100 |
| Table 5.10 | Correlation between national item parameter estimates for Hungarian versions..... | 100 |
| Table 5.11 | Correlation between national item parameter estimates for Italian versions..... | 100 |
| Table 5.12 | Correlation between national item parameter estimates for Portuguese versions..... | 100 |
| Table 5.13 | Correlation between national item parameter estimates for Russian versions..... | 100 |
| Table 5.14 | Correlation between national item parameter estimates for Spanish versions | 100 |
| Table 5.15 | Correlation between national item parameter estimates for Swedish versions | 100 |
| Table 5.16 | Correlation between national item parameter estimates within countries..... | 101 |
| Table 5.17 | Variance estimate..... | 102 |
| Table 5.18 | Variance estimates | 103 |
| Table 6.1 | Design for the multiple coding of science and mathematics..... | 118 |
| Table 6.2 | Design for the multiple coding of reading..... | 118 |
| Table 8.1 | Non-response classes | 133 |
| Table 9.1 | Deviation contrast coding scheme | 154 |
| Table 10.1 | Double entry discrepancies per country: field trial data..... | 169 |
| Table 11.1 | Sampling and coverage rates..... | 178 |
| Table 11.2 | School response rates before replacement..... | 182 |
| Table 11.3 | School response rates after replacement..... | 184 |
| Table 11.4 | Student response rates after replacement..... | 185 |
| Table 11.5 | Standard errors for the PISA 2006 combined science scale | 189 |
| Table 11.6 | Design effect 1 by country, by domain and cycle..... | 193 |
| Table 11.7 | Effective sample size 1 by country, by domain and cycle..... | 194 |
| Table 11.8 | Design effect 2 by country, by domain and cycle..... | 195 |
| Table 11.9 | Effective sample size 2 by country, by domain and cycle..... | 196 |
| Table 11.10 | Design effect 3 by country, by domain and by cycle..... | 197 |



| | | |
|-------------|--|-----|
| Table 11.11 | Effective sample size 3 by country, by domain and cycle | 198 |
| Table 11.12 | Design effect 4 by country, by domain and cycle..... | 199 |
| Table 11.13 | Effective sample size 4 by country, by domain and cycle | 200 |
| Table 11.14 | Design effect 5 by country, by domain and cycle..... | 201 |
| Table 11.15 | Effective sample size 5 by country, by domain and cycle | 202 |
| Table 11.16 | Median of the design effect 3 per cycle and per domain across the 35 countries that participated in every cycle..... | 203 |
| Table 11.17 | Median of the standard errors of the student performance mean estimate for each domain and PISA cycle for the 35 countries that participated in every cycle | 203 |
| Table 11.18 | Median of the number of participating schools for each domain and PISA cycle for the 35 countries that participated in every cycle..... | 204 |
| Table 11.19 | Median of the school variance estimate for each domain and PISA cycle for the 35 countries that participated in every cycle..... | 204 |
| Table 11.20 | Median of the intraclass correlation for each domain and PISA cycle for the 35 countries that participated in every cycle..... | 204 |
| Table 11.21 | Median of the within explicit strata intraclass correlation for each domain and PISA cycle for the 35 countries that participated in every cycle | 205 |
| Table 11.22 | Median of the percentages of school variances explained by explicit stratification variables, for each domain and PISA cycle for the 35 countries that participated in every cycle | 205 |
| <hr/> | | |
| Table 12.1 | Number of sampled student by country and booklet..... | 209 |
| Table 12.2 | Reliabilities of each of the four overall scales when scaled separately..... | 215 |
| Table 12.3 | Latent correlation between the five domains | 215 |
| Table 12.4 | Latent correlation between science scales | 215 |
| Table 12.5 | Items deleted at the national level | 216 |
| Table 12.6 | Final reliability of the PISA scales | 216 |
| Table 12.7 | National reliabilities for the main domains..... | 217 |
| Table 12.8 | National reliabilities for the science subscales..... | 218 |
| Table 12.9 | Average number of not-reached items and missing items by booklet..... | 219 |
| Table 12.10 | Average number of not-reached items and missing items by country..... | 220 |
| Table 12.11 | Distribution of not-reached items by booklet | 221 |
| Table 12.12 | Estimated booklet effects on the PISA scale..... | 221 |
| Table 12.13 | Estimated booklet effects in logits | 221 |
| Table 12.14 | Variance in mathematics booklet means | 222 |
| Table 12.15 | Variance in reading booklet means..... | 224 |
| Table 12.16 | Variance in science booklet means..... | 226 |
| Table 12.17 | Variance in interest booklet means | 228 |
| Table 12.18 | Variance in support booklet means..... | 230 |
| Table 12.19 | Summary of PISA cognitive reporting scales | 233 |
| Table 12.20 | Linkage types among PISA domains 2000-2006 | 235 |
| Table 12.21 | Number of unique item minutes for each domain for each PISA assessments..... | 237 |
| Table 12.22 | Numbers of link items between successive PISA assessments..... | 237 |
| Table 12.23 | Per cent correct for reading link items in PISA 2000 and PISA 2003 | 238 |
| Table 12.24 | Per cent correct for reading link items in PISA 2003 and PISA 2006 | 239 |
| Table 12.25 | Per cent correct for mathematics link items in PISA 2003 and PISA 2006 | 241 |

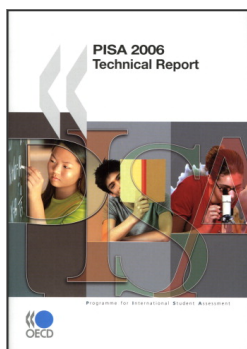
| | | |
|-------------|--|-----|
| Table 12.26 | Per cent correct for science link items in PISA 2000 and PISA 2003 | 243 |
| Table 12.27 | Per cent correct for science link items in PISA 2003 and PISA 2006 | 245 |
| Table 12.28 | Link error estimates | 247 |
| <hr/> | | |
| Table 13.1 | Variance components for mathematics..... | 255 |
| Table 13.2 | Variance components for science | 256 |
| Table 13.3 | Variance components for reading..... | 257 |
| Table 13.4 | Generalisability estimates for mathematics..... | 258 |
| Table 13.5 | Generalisability estimates for science | 259 |
| Table 13.6 | Generalisability estimates for reading | 260 |
| Table 13.7 | Examples of flagged cases | 263 |
| Table 13.8 | Count of analysis groups showing potential bias, by domain..... | 264 |
| Table 13.9 | Comparison of codes assigned by verifier and adjudicator | 265 |
| Table 13.10 | Outcomes of ICR analysis part 1 | 265 |
| Table 13.11 | ICR outcomes by country and domain | 266 |
| <hr/> | | |
| Table 15.1 | Scientific literacy performance band definitions on the PISA scale | 293 |
| <hr/> | | |
| Table 16.1 | ISCO major group white-collar/blue-collar classification | 306 |
| Table 16.2 | ISCO occupation categories classified as science-related occupations | 307 |
| Table 16.3 | OECD means and standard deviations of WL estimates | 311 |
| Table 16.4 | Median, minimum and maximum percentages of between-school variance for student-level indices across countries..... | 313 |
| Table 16.5 | Household possessions and home background indices..... | 316 |
| Table 16.6 | Scale reliabilities for home possession indices in OECD countries | 317 |
| Table 16.7 | Scale reliabilities for home possession indices in partner countries/economies | 318 |
| Table 16.8 | Item parameters for interest in science learning (INTSCIE)..... | 318 |
| Table 16.9 | Item parameters for enjoyment of science (JOYSCIE) | 319 |
| Table 16.10 | Model fit and estimated latent correlations for interest in and enjoyment of science learning..... | 319 |
| Table 16.11 | Scale reliabilities for interest in and enjoyment of science learning..... | 320 |
| Table 16.12 | Item parameters for instrumental motivation to learn science (INSTSCIE)..... | 320 |
| Table 16.13 | Item parameters for future-oriented science motivation (SCIEFUT)..... | 321 |
| Table 16.14 | Model fit and estimated latent correlations for motivation to learn science | 321 |
| Table 16.15 | Scale reliabilities for instrumental and future-oriented science motivation..... | 322 |
| Table 16.16 | Item parameters for science self-efficacy (SCIEEFF)..... | 322 |
| Table 16.17 | Item parameters for science self-concept (SCSCIE)..... | 323 |
| Table 16.18 | Model fit and estimated latent correlations for science self-efficacy and science self-concept..... | 323 |
| Table 16.19 | Scale reliabilities for science self-efficacy and science self-concept..... | 324 |
| Table 16.20 | Item parameters for general value of science (GENSCIE)..... | 324 |
| Table 16.21 | Item parameters for personal value of science (PERSCIE)..... | 325 |
| Table 16.22 | Model fit and estimated latent correlations for general and personal value of science..... | 325 |
| Table 16.23 | Scale reliabilities for general and personal value of science..... | 326 |
| Table 16.24 | Item parameters for science activities (SCIEACT) | 326 |



| | | |
|-------------|--|-----|
| Table 16.25 | Scale reliabilities for the science activities index | 327 |
| Table 16.26 | Item parameters for awareness of environmental issues (ENVAWARE) | 327 |
| Table 16.27 | Item parameters for perception of environmental issues (ENVPERC) | 328 |
| Table 16.28 | Item parameters for environmental optimism (ENVOPT) | 328 |
| Table 16.29 | Item parameters for responsibility for sustainable development (RESPDEV) | 328 |
| Table 16.30 | Model fit environment-related constructs | 329 |
| Table 16.31 | Estimated latent correlations for environment-related constructs | 329 |
| Table 16.32 | Scale reliabilities for environment-related scales in OECD countries | 330 |
| Table 16.33 | Scale reliabilities for environment-related scales in non-OECD countries | 330 |
| Table 16.34 | Item parameters for school preparation for science career (CARPREP) | 331 |
| Table 16.35 | Item parameters for student information on science careers (CARINFO) | 331 |
| Table 16.36 | Model fit and estimated latent correlations for science career preparation indices | 332 |
| Table 16.37 | Scale reliabilities for science career preparation indices | 332 |
| Table 16.38 | Item parameters for science teaching: interaction (SCINTACT) | 333 |
| Table 16.39 | Item parameters for science teaching: hands-on activities (SCHANDS) | 333 |
| Table 16.40 | Item parameters for science teaching: student investigations (SCINVEST) | 333 |
| Table 16.41 | Item parameters for science teaching: focus on models or applications (SCAPPLY) | 334 |
| Table 16.42 | Model fit for CFA with science teaching and learning | 334 |
| Table 16.43 | Estimated latent correlations for constructs related to science teaching and learning | 335 |
| Table 16.44 | Scale reliabilities for scales to science teaching and learning in OECD countries | 336 |
| Table 16.45 | Scale reliabilities for scales to science teaching and learning in partner countries/economies | 336 |
| Table 16.46 | Item parameters for ICT Internet/entertainment use (INTUSE) | 337 |
| Table 16.47 | Item parameters for ICT program/software use (PRGUSE) | 337 |
| Table 16.48 | Item parameters for ICT self-confidence in Internet tasks (INTCONF) | 337 |
| Table 16.49 | Item parameters for ICT self-confidence in high-level ICT tasks (HIGHCONF) | 338 |
| Table 16.50 | Model fit for CFA with ICT familiarity items | 338 |
| Table 16.51 | Estimated latent correlations for constructs related to ICT familiarity | 339 |
| Table 16.52 | Scale reliabilities for ICT familiarity scales | 339 |
| Table 16.53 | Item parameters for teacher shortage (TCSHORT) | 340 |
| Table 16.54 | Item parameters for quality of educational resources (SCMATEDU) | 340 |
| Table 16.55 | Item parameters for school activities to promote the learning of science (SCIPROM) | 341 |
| Table 16.56 | Item parameters for school activities for learning environmental topics (ENVLEARN) | 341 |
| Table 16.57 | Scale reliabilities for school-level scales in OECD countries | 341 |
| Table 16.58 | Scale reliabilities for environment-related scales in partner countries/economies | 342 |
| Table 16.59 | Item parameters for science activities at age 10 (PQSCIACT) | 343 |
| Table 16.60 | Item parameters for parent's perception of school quality (PQSCHOOL) | 343 |
| Table 16.61 | Item parameters for parent's views on importance of science (PQSCIMP) | 343 |
| Table 16.62 | Item parameters for parent's reports on science career motivation (PQSCCAR) | 344 |
| Table 16.63 | Item parameters for parent's view on general value of science (PQGENSCI) | 344 |
| Table 16.64 | Item parameters for parent's view on personal value of science (PQPERSCI) | 344 |
| Table 16.65 | Item parameters for parent's perception of environmental issues (PQENPERC) | 345 |
| Table 16.66 | Item parameters for parent's environmental optimism (PQENVOPT) | 345 |



| | | |
|-------------|--|-----|
| Table 16.67 | Scale reliabilities for parent questionnaire scales..... | 345 |
| Table 16.68 | Factor loadings and internal consistency of ESCS 2006 in OECD countries..... | 347 |
| Table 16.69 | Factor loadings and internal consistency of ESCS 2006 in partner countries/economies..... | 348 |
| <hr/> | | |
| Table 17.1 | Student-level latent correlations between mathematics, reading, science, embedded interest and embedded support..... | 354 |
| Table 17.2 | Summary of the IRT scaling results across countries..... | 355 |
| Table 17.3 | Gender DIF table for embedded attitude items..... | 357 |
| Table 17.4 | Correlation amongst attitudinal scales, performance scales and HISEI..... | 358 |
| Table 17.5 | Correlations for science scale..... | 359 |
| Table 17.6 | Loadings of the achievement, interest and support variables on three varimax rotated components..... | 360 |
| Table 17.7 | Correlation between embedded attitude scales and questionnaire attitude scales..... | 361 |
| Table 17.8 | Rank order correlation five test domains, questionnaire attitude scales and HISEI..... | 362 |
| Table 17.9 | Intra-class correlation (rho)..... | 362 |
| <hr/> | | |
| Table A1.1 | 2006 Main study reading item classification..... | 380 |
| Table A1.2 | 2006 Main study mathematics item classification..... | 381 |
| Table A1.3 | 2006 Main study science item classification (cognitive)..... | 383 |
| Table A1.4 | 2006 Main study science embedded item classification (interest in learning science topics)..... | 387 |
| Table A1.5 | 2006 Main study science embedded item classification (support for scientific enquiry)..... | 388 |
| <hr/> | | |
| Table A2.1 | 2006 Main study contrast coding used in conditioning for the student questionnaire variables..... | 389 |
| Table A2.2 | 2006 Main study contrast coding used in conditioning for the ICT questionnaire variables..... | 396 |
| Table A2.3 | 2006 Main study contrast coding used in conditioning for the parent questionnaire variables and other variables..... | 397 |
| <hr/> | | |
| Table A3.1 | Standard errors of the student performance mean estimate by country, by domain and cycle..... | 399 |
| Table A3.2 | Sample sizes by country and cycle..... | 400 |
| Table A3.3 | School variance estimate by country, by domain and cycle..... | 401 |
| Table A3.4 | Intraclass correlation by country, by domain and cycle..... | 402 |
| Table A3.5 | Within explicit strata intraclass correlation by country, by domain and cycle..... | 403 |
| Table A3.6 | Percentages of school variance explained by explicit stratification variables, by domain and cycle..... | 404 |
| <hr/> | | |
| Table A4.1 | Student questionnaire..... | 405 |
| Table A4.2 | ICT familiarity questionnaire..... | 407 |
| Table A4.3 | School questionnaire..... | 408 |
| <hr/> | | |
| Table A5.1 | Mapping of ISCED to accumulated years of education..... | 411 |
| <hr/> | | |
| Table A6.1 | National household possession items..... | 412 |
| <hr/> | | |
| Table A7.1 | Exploratory and confirmatory factor analyses (EFA and CFA) for the embedded items..... | 414 |



From:
PISA 2006 Technical Report

Access the complete publication at:
<https://doi.org/10.1787/9789264048096-en>

Please cite this chapter as:

OECD (2009), "The development of the PISA context questionnaires", in *PISA 2006 Technical Report*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264048096-4-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.